

Erraticity of Rapidity Gaps

Rudolph C. Hwa and Qing-hui Zhang

Institute of Theoretical Science and Department of Physics
University of Oregon, Eugene, OR 97403-5203

Abstract

The use of rapidity gaps is proposed as a measure of the spatial pattern of an event. When the event multiplicity is low, the gaps between neighboring particles carry far more information about an event than multiplicity spikes, which may occur very rarely. Two moments of the gap distribution are suggested for characterizing an event. The fluctuations of those moments from event to event are then quantified by an entropy-like measure, which serves to describe erraticity. We use ECOMB to simulate the exclusive rapidity distribution of each event, from which the erraticity measures are calculated. The dependences of those measures on the order of q of the moments provide single-parameter characterizations of erraticity.

1 Introduction

To study the properties of event-to-event fluctuations in multiparticle production, it is necessary to have an effective measure of the characteristics of the final state of an event. The totality of all the momenta of the produced particles constitutes a pattern. A useful measure of a pattern should not contain too much details, but enough to capture the essence that is likely to fluctuate from event-to-event. In previous papers we have used the normalized factorial moments, F_q , as a measure for studying chaos in QCD jets [1], in classical nonlinear dynamics [2], and erraticity in soft production of particles [3]. We now consider two different moments in order to improve the analysis in problems where the use of F_q is less effective.

The horizontal factorial moments register the multiplicity fluctuation from bin to bin in an event. However, when the event multiplicity is low, and the bin size small, most bins have only one particle per bin, and the factorial moments fail to provide a good characterization of the event pattern. To overcome that deficiency, we shift our emphasis from bin multiplicities to rapidity gaps. It is intuitively obvious that the two quantities are complementary: the former counts how many particles fall into the same bin, while the latter measures how far apart neighboring particles are. Clearly, the former works better when there are many particles in an event, while the latter is more suitable when there are few particles.

The search for a good measure of event-to-event fluctuations can be carried out only if we have an event generator that can be used for the exploration. To that end we shall use ECOMB [4], which simulates soft production processes in hadronic collision. It is based on a reasonable modeling of the many-body dynamics in which the partons undergo successive

color mutation before hadronization. It is the only model capable of generating factorial moments that agree with the intermittency data of NA22 [5]. However, there is still freedom in the model for further adjustment. The aim of this paper is not to test ECOMB or to improve it. Despite its imperfections, it can nevertheless simulate events with sufficient dynamical fluctuations that deviate significantly from statistical fluctuations. That capability is what we utilize in our search for the desired measure. For that reason it is unnecessary for us to review here the dynamical content of ECOMB. After the experimental data are analyzed and the proposed measure determined, we can then return to the problem of modeling soft interaction when the new erraticity data will provide the guidance needed for an upgrading of ECOMB.

2 The Problem

Let us start by reviewing the factorial moments for multiparticle production [6]. They are defined (for the q th order) by

$$f_q = \langle n(n-1) \cdots (n-q+1) \rangle \quad , \quad (1)$$

where n is the multiplicity in a bin. Originally, the average in (1) is performed over all events for a fixed bin, which we now call vertical average. Later, the horizontal average is considered for a fixed event, where n in (1) is averaged over all bins. In either case if the probability distribution P_n of n can be expressed as a convolution of the dynamical distribution $D(\nu)$ and the statistical (Poissonian) distribution, i.e.,

$$P_n = \int d\nu \frac{1}{n!} \nu^n e^{-\nu} D(\nu) \quad , \quad (2)$$

then one obtains [6]

$$f_q = \sum_{n=q}^{\infty} \frac{n!}{(n-q)!} P_n = \int d\nu \nu^q D(\nu) \quad . \quad (3)$$

Since it is a simple moment of $D(\nu)$, the statistical fluctuation is regarded as having been filtered out by f_q .

The above procedure of eliminating the statistical fluctuation fails either when the sum in (3) does not extend to infinity, or if that fluctuation cannot be represented by a Poissonian distribution as in (2). Both of these circumstances occur for horizontal analyses of low multiplicity events. There is nothing wrong with calculating f_q according to (1) for such events. The question is what one can use f_q for.

In [3] the horizontal normalized factorial moments $F_q = f_q/f_1^q$ are used to characterize the spatial pattern of an event. Such a characterization clearly cannot convey all the details of an event; indeed, extensive details using many variables are not desirable for the quantification of event-to-event fluctuations. It is evident from (1) that only bins with $n \geq q$ can contribute to f_q , but the positions of the contributing bins have no effect on f_q . That deficiency is unimportant when many bins contribute. However, when the event multiplicity N is low and the number of bins M is high, so that the average bin multiplicity $\bar{n} = N/M \ll 1$, then

it is only by large fluctuations that a bin may have $n \geq q$, whether they are dynamical or statistical in nature. Since f_q is insensitive to where the few contributing bins are located, there is very little information about an event that is registered in F_q . In [7] it is shown that in the framework of a simple model F_q are dominated by statistical fluctuations when N is small, but they reveal the dynamical fluctuations when N is large.

The aim of this paper is to find an alternative to F_q that can effectively characterize the spatial pattern of an event, even when the event multiplicity is low.

3 The Solution

From (1) we see that f_q receives a contribution from a bin in which $n \geq q$, but ignores where it is located. In other words f_q is sensitive to the local height of the rapidity distribution in an event, not to the spatial arrangement in rapidity. When N is low and M is high, many bins are empty. To have a bin with $n \geq q$ means that even more bins than average would have to be empty. It then seems clear that the complementary information accompanying rapidity spikes is the rapidity gaps. When N is high, rapidity gaps are generally not very informative; however, when N is low, they characterize an event better than counting spikes. In the following we shall develop two methods based on measuring the rapidity gaps.

Since particle momenta can be measured accurately, there is no need to consider discrete bins in the rapidity space. Thus we shall work in the continuum. Moreover, the advantage of working in the cumulative variable X has long been recognized [8, 9], and we shall continue to use the X variable, as in [4] (though not explicitly stated there in the first reference). The definition of X is

$$X(y) = \int_{y_{\min}}^y \rho(y') dy' / \int_{y_{\min}}^{y_{\max}} \rho(y') dy' \quad , \quad (4)$$

where $\rho(y')$ is the single-particle inclusive rapidity distribution and $y_{\min(\max)}$ is the minimum (maximum) value of y . Thus the accessible range of y is mapped to X between 0 and 1, and the density of particles in X , dn/dX , is uniform.

Consider an event with N particles, labeled by $i = 1, \dots, N$, located in the X space at X_i , ordered from the left to the right. Let us now define the distance between neighboring particles by

$$x_i = X_{i+1} - X_i, \quad i = 0, \dots, N \quad , \quad (5)$$

with $X_0 = 0$ and $X_{i+1} = 1$ being the boundaries of the X space. Every event e is thus characterized by a set S_e of $N + 1$ number: $S_e = \{x_i | i = 0, \dots, N\}$, which clearly satisfy

$$\sum_{i=0}^N x_i = 1 \quad . \quad (6)$$

We refer to these numbers loosely as “rapidity” gaps.

For any given event S_e contains more information than P_n , which is the bin-multiplicity distribution for that event; in fact, P_n can be determined from S_e , but not in reverse. To study the fluctuation of S_e from event-to-event is the most that one can do; indeed, too

much information is conveyed by S_e . For economy and efficiency in codifying the information we consider moments of x_i that emphasize large rapidity gaps. As mentioned earlier, concomitant to the clustering of particles that results in spikes in the rapidity distribution is the existence of large gaps. Thus moments that emphasize large x_i convey similar information about an event as do the moments that emphasize the high- n tail of P_n . However, the factorial moments of P_n suffer the defects discussed in Sec. 2 that are absent in the moments of x_i . The issue of statistical fluctuations has to be addressed separately.

Let us then define for each event

$$G_q = \frac{1}{N+1} \sum_{i=0}^N x_i^q \quad , \quad (7)$$

Despite the similarity in notation, these moments bear no relationship to the G -moments considered earlier [10]. It is clear from (6) and (7) that

$$G_0 = 1 \quad \text{and} \quad G_1 = \frac{1}{N+1} \quad . \quad (8)$$

At higher q , G_q are progressively smaller, but are increasingly more dominated by the large x_i components in S_e . A set of G_q for q ranging up to 5 or 6 is sufficient to characterize an event, better than S_e itself in the sense that G_q can be compared from event-to-event, whereas S_e cannot be so compared due to the fluctuations in N .

If we define the gap distribution by

$$g(x) = \frac{1}{N+1} \sum_{i=0}^N \delta(x - x_i) \quad , \quad (9)$$

then the G moments are

$$G_q = \int_0^1 dx x^q g(x) \quad . \quad (10)$$

This form may become more convenient in some situations.

Since G_q fluctuates from event to event, we can determine a distribution $P(G_q)$ of G_q after many events. It is the shape of $P(G_q)$ that characterizes the nature of the event-to-event fluctuations of the gap distribution, and therefore of the spatial pattern of an event. Again, we can describe $P(G_q)$ by its moments

$$C_{p,q} = \frac{1}{\mathcal{N}} \sum_{e=1}^{\mathcal{N}} (G_q^e)^p = \int dG_q G_q^p P(G_q) \quad , \quad (11)$$

where e labels an event and \mathcal{N} is the total number of events. Since we need not consider bins in x , G_q is a number for each event without statistical error. Thus calculating the p th moment does not compound statistical errors. Although one can consider a range of p moments, we shall focus only on the derivative at $p = 1$ in the following.

Since $C_{1,q} = \langle G_q \rangle$ is the mean that gives no information on the degree of fluctuation, the derivative at $p = 1$ convey the broadest information on $P(G_q)$. We have

$$s_q = -\frac{d}{dp} C_{p,q} \big|_{p=1} = -\langle G_q \ln G_q \rangle \quad , \quad (12)$$

where $\langle \dots \rangle$ stands for average over all events. The quantities s_q are our new measures of erraticity in terms of rapidity gaps. Since G_q is not a probability distribution, s_q is not an entropy function, despite its appearance.

Unlike the factorial moments, G_q does not filter out statistical fluctuations. At low multiplicities F_q fails to be effective in that filtering anyway, as discussed in Sec. 2, so it is at no great loss to consider G_q . However, we can have an estimate of how much s_q stands out above the statistical fluctuation by first calculating

$$s_q^{st} = - \left\langle G_q^{st} \ln G_q^{st} \right\rangle \quad , \quad (13)$$

where G_q^{st} is determined from (10) by using only the statistical distribution of the gaps, $g^{st}(x)$, i.e., when all N particles in an event are distributed randomly in X space. Then we take the ratio

$$S_q = s_q / s_q^{st} \quad , \quad (14)$$

and examine how much S_q deviates from 1. S_q will be the first erraticity measure that we shall calculate in the next section.

Since our interest is in the deviation of G_q from $\langle G_q \rangle$, a measure of that deviation is

$$\tilde{s}_q = - \left\langle \frac{G_q}{\langle G_q \rangle} \ln \frac{G_q}{\langle G_q \rangle} \right\rangle \quad , \quad (15)$$

which clearly would be zero if G_q never deviates from $\langle G_q \rangle$. We can further normalize \tilde{s}_q^{st} by the statistical-only contribution \tilde{s}_q^{st} and define

$$\tilde{S}_q = \tilde{s}_q / \tilde{s}_q^{st} \quad . \quad (16)$$

Whether this is a better quantity to represent erraticity will be examined quantitatively in the next section.

The moments G_q are not the only ones that can characterize the rapidity-gap distribution. In fact, since $x_i < 1$, G_q are usually $\ll 1$, and the statistical errors on S_q and \tilde{S}_q turn out to be quite large, though not so large as to render the measures ineffective. We now consider a different type of moments that also emphasize the large gaps. Define for an event with N particles

$$H_q = \frac{1}{N+1} \sum_{i=0}^N (1-x_i)^{-q} \quad , \quad (17)$$

where x_i is as given in (5). These moments also receive dominant contribution from large x_i , as do G_q , but H_q can become $\gg 1$. In terms of $g(x)$ we have

$$H_q = \int_0^1 dx (1-x)^{-q} g(x) \quad , \quad (18)$$

where $g(x)$ must vanish sufficiently fast as $x \rightarrow 1$ to safeguard the integrability of (18).

We can substitute H_q for G_q in all of the foregoing considerations. In particular, we can define

$$\sigma_q = \langle H_q \ln H_q \rangle \quad , \quad (19)$$

$$\tilde{\sigma}_q = \left\langle \frac{H_q}{\langle H_q \rangle} \ln \frac{H_q}{\langle H_q \rangle} \right\rangle \quad , \quad (20)$$

$$\Sigma_q = \frac{\sigma_q}{\sigma_q^{st}} \quad \text{and} \quad \tilde{\Sigma}_q = \frac{\tilde{\sigma}_q}{\tilde{\sigma}_q^{st}} \quad (21)$$

as new measures of erraticity. The only nontrivial point to remark on concerns the event average.

For each event H_q depends implicitly on the event multiplicity N . If \mathcal{P}_n is the multiplicity distribution, then the average $\langle H_q \rangle$ is given by

$$\langle H_q \rangle = \sum_{N=q+1}^{\infty} H_q(N) \mathcal{P}_N \quad , \quad (22)$$

where $H_q(N)$ is the mean H_q after averaging over all events with N particles in each event. Note that the sum in (22) begins at $N = q + 1$, not 0. To see this subtle point, let us start with the statistical average for which we can make precise calculations. In the Appendix we show that the probability distribution $p_N(x)$ of the gap distance x , after sampling with sufficiently many events, each with N randomly distributed particles in the X space, is

$$p_N^{st}(x) = N (1 - x)^{N-1} \quad . \quad (23)$$

Thus it follows that

$$H_q^{st}(N) = \int_0^1 dx (1 - x)^{-q} p_N^{st}(x) = \frac{N}{N - q} \quad . \quad (24)$$

Evidently, N must be greater than q to ensure convergence. If for statistical calculation we require $n \geq q + 1$, then we make the same requirement for the general problem in (22), so that σ_q and σ_q^{st} in (21) are calculated on the same basis.

4 Results

We have applied ECOMB [4], upgraded by [11], to calculate the rapidity distribution for each event. From that we compute the gap distribution $g(x)$ in the X space. After simulating 10^6 events at $\sqrt{s} = 20$ GeV, our result for S_q is shown in Fig. 1. The error bars are determined by using the conventional method. The straight line in Fig. 1 is a linear fit of the central points. Evidently, the result indicates a power-law behavior in q for $q \geq 2$

$$S_q \propto q^\alpha \quad , \quad \alpha = 0.156 \quad . \quad (25)$$

The fact that S_q deviates unambiguously from 1 implies that it is a statistically significant measure of erraticity in multiparticle production. At $\sqrt{s} = 20$ GeV the average charge multiplicity is only 8.5, which is low enough to cause problems for the factorial moments F_q , but our use of the gap moments G_q evidently encounters no similar difficulty.

We next consider \tilde{S}_q defined in (15) and (16). The result is that \tilde{S}_q is nearly independent of q , as shown in Fig. 2. More precisely, we obtain

$$\tilde{S}_q = 0.96 \pm 0.03 \quad . \quad (26)$$

We regard this result as indicative of the inadequacy of \tilde{S}_q as a measure of erraticity, since \tilde{S}_q is almost consistent with 1.

Turning to the H_q moments, we show in Fig. 3 in semilog plot the dependence of Σ_q on q . Evidently, a very good linear fit is obtained, yielding

$$\Sigma_q \propto e^{\beta q} \quad , \quad \beta = 0.28 \quad . \quad (27)$$

In the same figure we also show $\tilde{\Sigma}_q$. Although the error bars are larger, an exponential behavior

$$\tilde{\Sigma}_q \propto e^{\tilde{\beta} q} \quad , \quad \tilde{\beta} = 0.25 \quad , \quad (28)$$

can nevertheless be identified. Note that $\tilde{\Sigma}_q$ is much farther from 1 than \tilde{S}_q . Since Σ_q has less statistical error than $\tilde{\Sigma}_q$, it is more preferred. Hereafter we shall discard \tilde{S}_q and $\tilde{\Sigma}_q$ from any further consideration.

We now examine the dependence on c.m. energy. The higher multiplicities at higher s will decrease the average gap $\langle x \rangle$ and the corresponding moments G_q and H_q . We calculate the effects on S_q and Σ_q at $\sqrt{s} = 200$ GeV. The results are shown in Figs. 4 and 5, where the values for $\sqrt{s} = 20$ GeV are reproduced for comparison. The power law (25) and the exponential behavior (27) persist; the corresponding parameters are

$$\alpha = 0.133 \quad , \quad \beta = 0.108 \quad \text{at} \quad \sqrt{s} = 200 \text{ GeV} \quad . \quad (29)$$

Whereas α has changed little, β has decreased significantly. The variability of β makes it a more sensitive measure of erraticity, although the stability of α may nevertheless be interesting and useful. Only the analysis of the experimental data will reveal which one between S_q and Σ_q is better in quantifying erraticity. It can also turn out that both are good.

5 Conclusion

We have proposed the moments G_q and H_q as measures of spatial patterns in terms of rapidity gaps. We then showed that the entropy-like quantities S_q and Σ_q deviate sufficiently from 1 with small enough statistical errors to serve as effective measures of erraticity, i.e., event-to-event fluctuations. In the framework of an soft hadronic interaction event generator ECOMB we have obtained the behaviors $S_q \propto q^\alpha$ and $\Sigma_q \propto e^{\beta q}$. The precise forms of these results are unimportant from the point of view of the search for an experimental measure to quantify

erraticity. We offer both S_q and Σ_q as our findings. On the other hand, from the point of view of using erraticity to test event generators, then the forms of our results for S_q and Σ_q are pertinent, and the values $\alpha = 0.156$ and $\beta = 0.28$ are useful for comparison with the soft production data. Analysis of the data, especially those of NA22, to determine S_q and Σ_q is therefore urged. The experimental values of α and β will either eliminate wrong models or provide crucial guidance to the improvement of the correct models.

The extension of this approach to other collision processes is obviously the next step. For heavy-ion collisions the multiplicities will be too high for any interesting study in rapidity gaps, unless one focuses on more rarely produced particles, such as J/ψ . At RHIC when only pp collisions are studied, our result for $\sqrt{s} = 200$ GeV can be tested. For nuclear collisions very narrow Δp_T selection must be made to render the rapidity gap analysis meaningful.

A natural direction of generalization is, of course, to higher dimensional analysis. One-dimensional gaps should be generalized to two-dimensional voids, which is more difficult to define if the use of bins is to be avoided. When a good measure is found, not only can it be employed as an alternative to the multiplicity analysis in the lego plot, useful application can no doubt be found also in the study of galactic clustering in astrophysics. Finally, in view of the abundance of experimental data and the variety of event generators for e^+e^- annihilation, a generalization to the multidimensional variables suitable for such problems will be a fruitful direction to pursue.

Acknowledgment

We are grateful to Dr. Z. Cao for helpful communication at the beginning of this research project. This work was supported in part by U. S. Department of Energy under Grant No. DE-FG03-96ER40972.

Appendix

We derive in this Appendix the probability distribution of gaps in the purely statistical case.

The gap distribution $g(x)$ defined in (9) and (5) can be more elaborately written as $g_e(x; X_1, \dots, X_N)$ for the e th event with N particles located at X_1, \dots, X_N . It has the normalization

$$\int_0^1 dx g_e(x; X_1, \dots, X_N) = 1 \quad . \quad (30)$$

If the value of X_i is randomly selected in the interval $0 \leq X_i \leq 1$, then after a large number of events \mathcal{N} with N particles in each, probability distribution in x is

$$\begin{aligned} p_N^{st}(x) &= \frac{1}{\mathcal{N}} \sum_{e=1}^{\mathcal{N}} g_e(x; X_1, \dots, X_N) \\ &= N! \int_0^1 dX_1 \int_{X_1}^1 dX_2 \cdots \int_{X_{N-1}}^1 dX_N g_e(x; X_1, \dots, X_N) \quad , \end{aligned} \quad (31)$$

in which the primitive distributions of the individual X_i values that should appear inside the integral have been set equal to 1 due to the statistical nature of their occurrences.

Consider a specific gap $y = X_j - X_i$, where $j = i + 1$. Then for all the multiple integrals at and before X_i , we may reverse the order of integration and obtain

$$\int_0^1 dX_1 \int_{X_1}^1 dX_2 \cdots \int_{X_{i-1}}^1 dX_i = \int_0^1 dX_i \int_0^{X_i} dX_{i-1} \cdots \int_0^{X_2} dX_1 = \int_0^1 dX_i \frac{1}{(i-1)!} X_i^{i-1} . \quad (32)$$

For all the multiple integrals at and after X_j , we have

$$\int_{X_j}^1 dX_{j+1} \cdots \int_{X_{N-1}}^1 dX_N = \frac{1}{(N-j)!} (1 - X_j)^{N-j} \quad . \quad (33)$$

Substituting these and (5) and (9) into (31), we get

$$\begin{aligned} p_N^{st}(x) &= \frac{N!}{N+1} \sum_{i=0}^N \int_0^1 dX_i \int_0^{X_i} dX_j \frac{X_i^{i-1}}{(i-1)!} \frac{(1 - X_j)^{N-j}}{(N-j)!} \delta(x - X_j + X_i) \\ &= \frac{N!}{N+1} \sum_{i=0}^N \int_0^1 dy \int_0^{1-y} dX_i \frac{X_i^{i-1}}{(i-1)!} \frac{(1 - X_i + y)^{N-j}}{(N-j)!} \delta(x - y) \quad . \end{aligned} \quad (34)$$

The integral over X_i yields $(1-y)^{N-1}/(N-1)!$. Thus the final result is

$$p_N^{st}(x) = N(1-x)^{N-1} \quad . \quad (35)$$

This behavior is verified by numerical simulation.

References

- [1] Z. Cao and R. C. Hwa, Phys. Rev. Lett. **75**, 1268 (1995); Phys. Rev. D **53**, 6608 (1996); *ibid* **54**, 6674 (1996).
- [2] Z. Cao and R. C. Hwa, Phys. Rev. E **56**, 326 (1997).
- [3] Z. Cao and R. C. Hwa, hep-ph/9901256, Phys. Rev. D (to be published).
- [4] Z. Cao and R. C. Hwa, Phys. Rev. D **59**, 114023 (1999).
- [5] I. V. Ajineko *et al.* (NA22), Phys. Lett. B **222**, 306 (1989); **235**, 373 (1990).
- [6] A. Białas and R. Peschanski, Nucl. Phys. B **273**, 703 (1986); **308**, 857 (1988).
- [7] J. Fu, Y. Wu, and L. Liu, hep-ph/9903217
- [8] A. Białas and M. Gardzicki, Phys. Lett. B **252**, 483 (1990)
- [9] E. A. DeWolf, I. M. Dremin, and W. Kittel, Phys. Rep. **270**, 1 (1996).
- [10] R. C. Hwa, Phys. Rev. D **41**, 1456 (1990); I. Derado, R. C. Hwa, G. Jancso, and N. Schmitz, Phys. Lett. B **283**, 151 (1992).
- [11] R. C. Hwa and Y. Wu, Phys. Rev. D **60**, 097501 (1999).

Figure Captions

Fig. 1 $\ell n S_q$ vs q as determined by ECOMB. The solid line is the best fit of the central points.

Fig. 2 \tilde{S}_q vs q with the same comments as in Fig. 1.

Fig. 3 $\ell n \Sigma_q$ and $\ell n \tilde{\Sigma}_q$ vs q with the same comments as in Fig. 1.

Fig. 4 $\ell n S_q$ vs q at two different energies.

Fig. 5 $\ell n \Sigma_q$ vs q at two different energies.

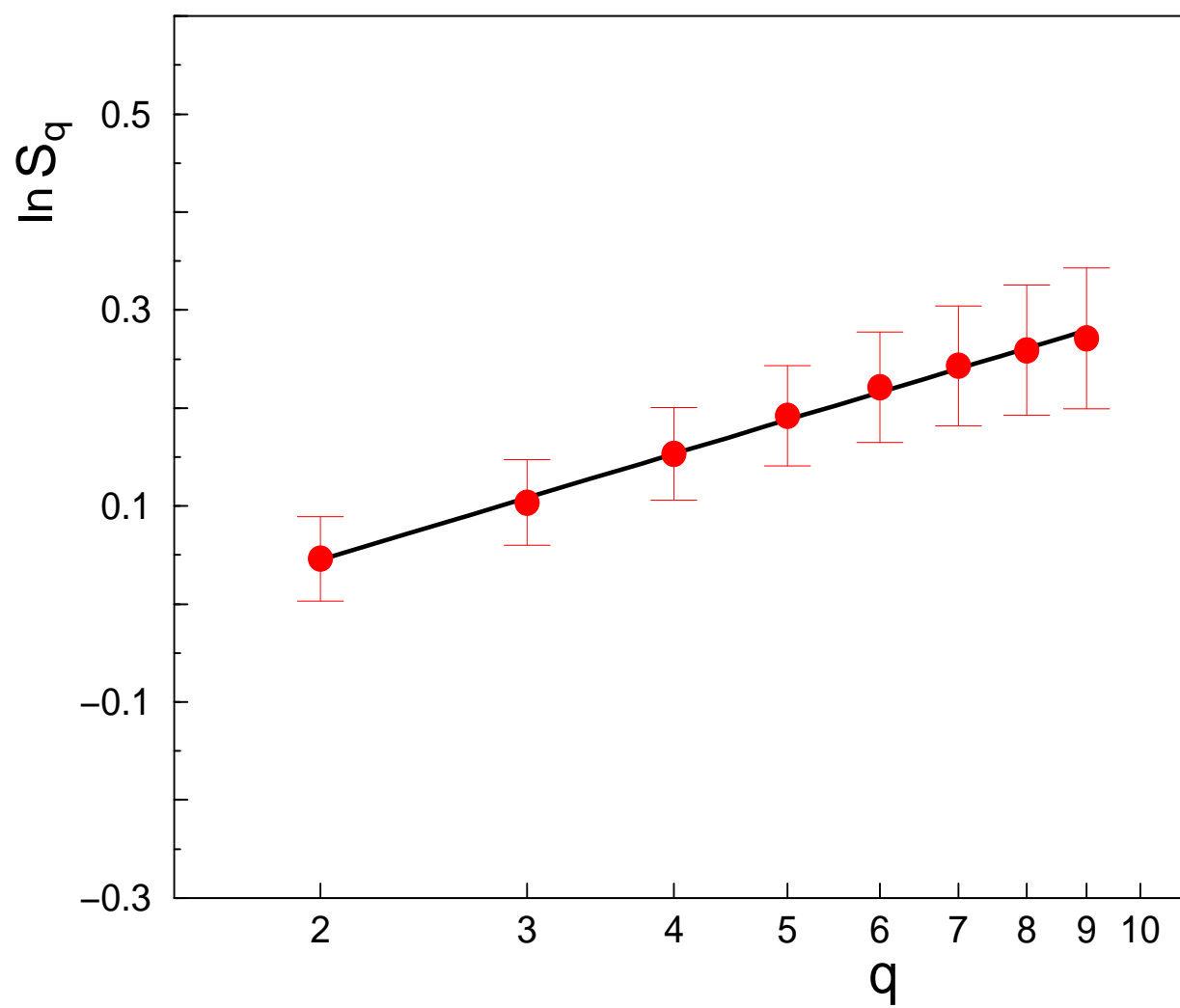


Fig.1

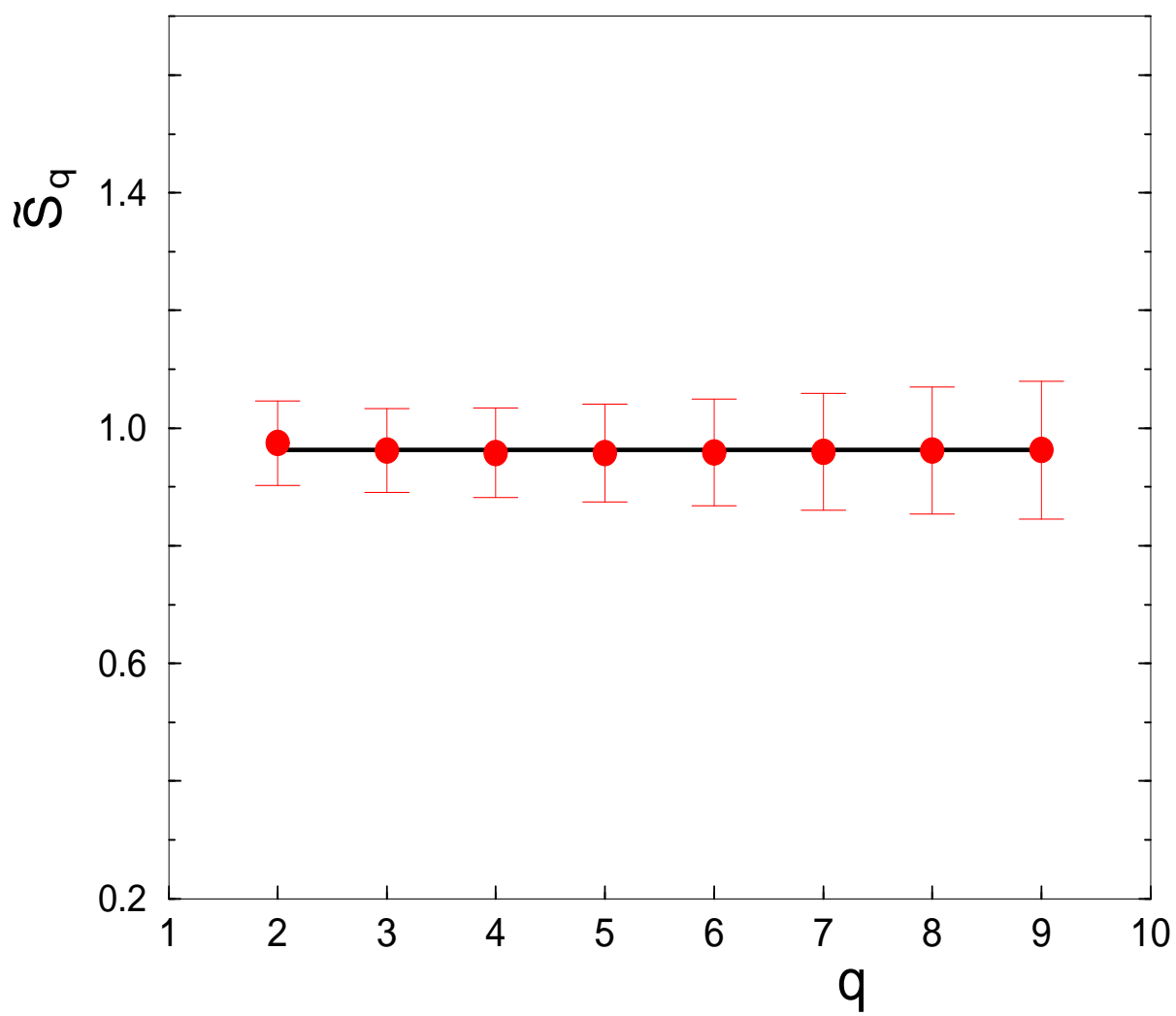


Fig.2

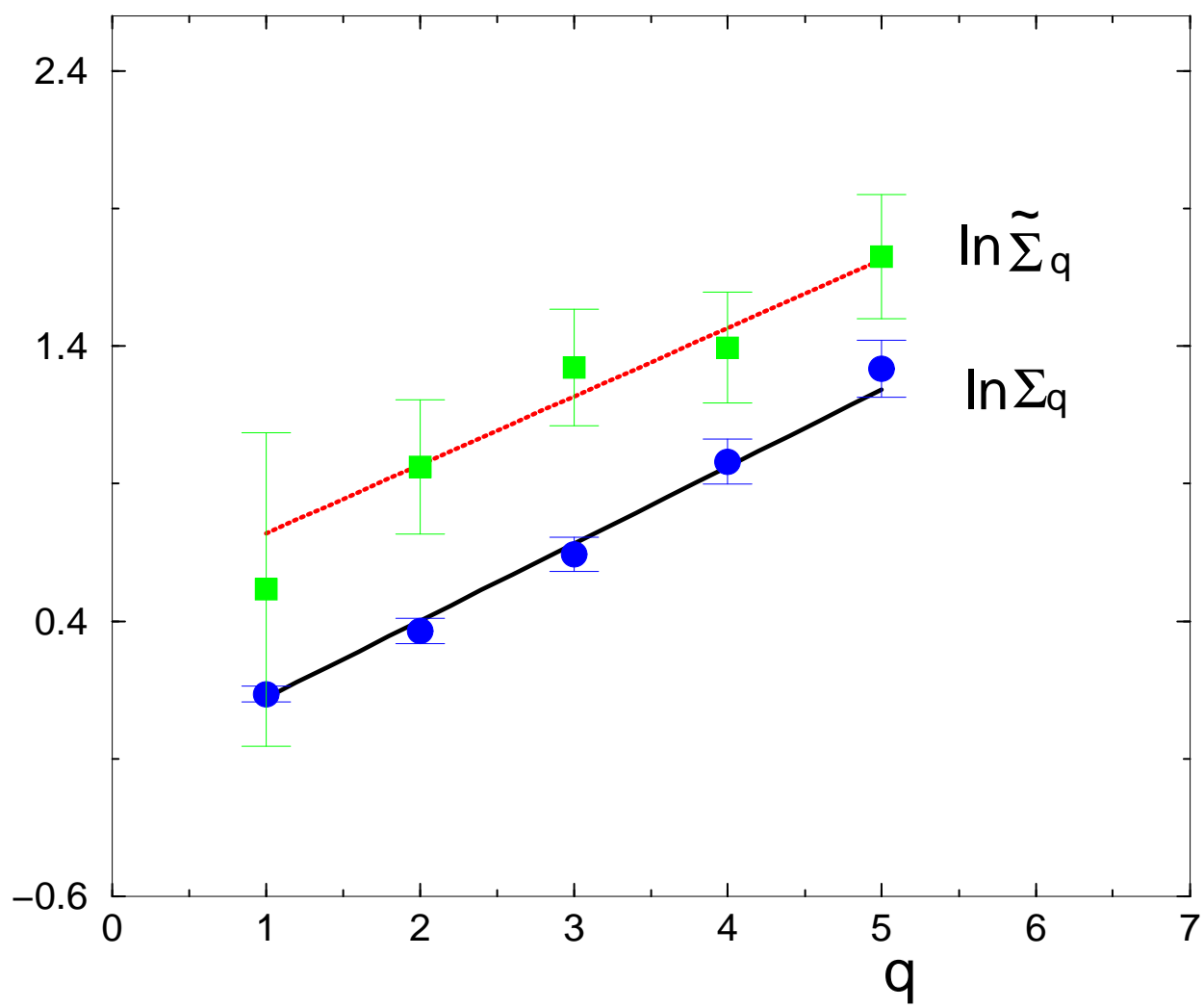


Fig.3

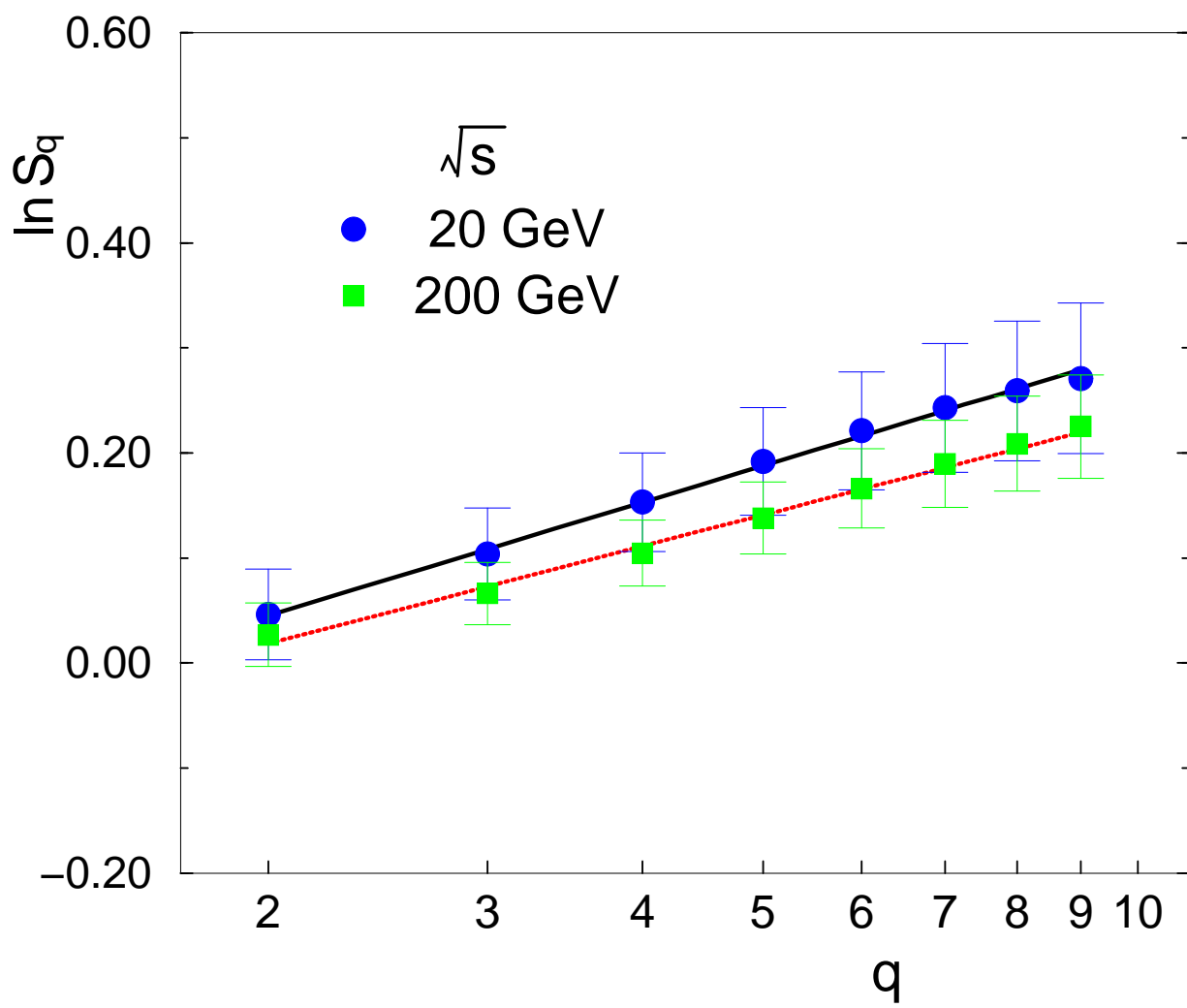


Fig.4

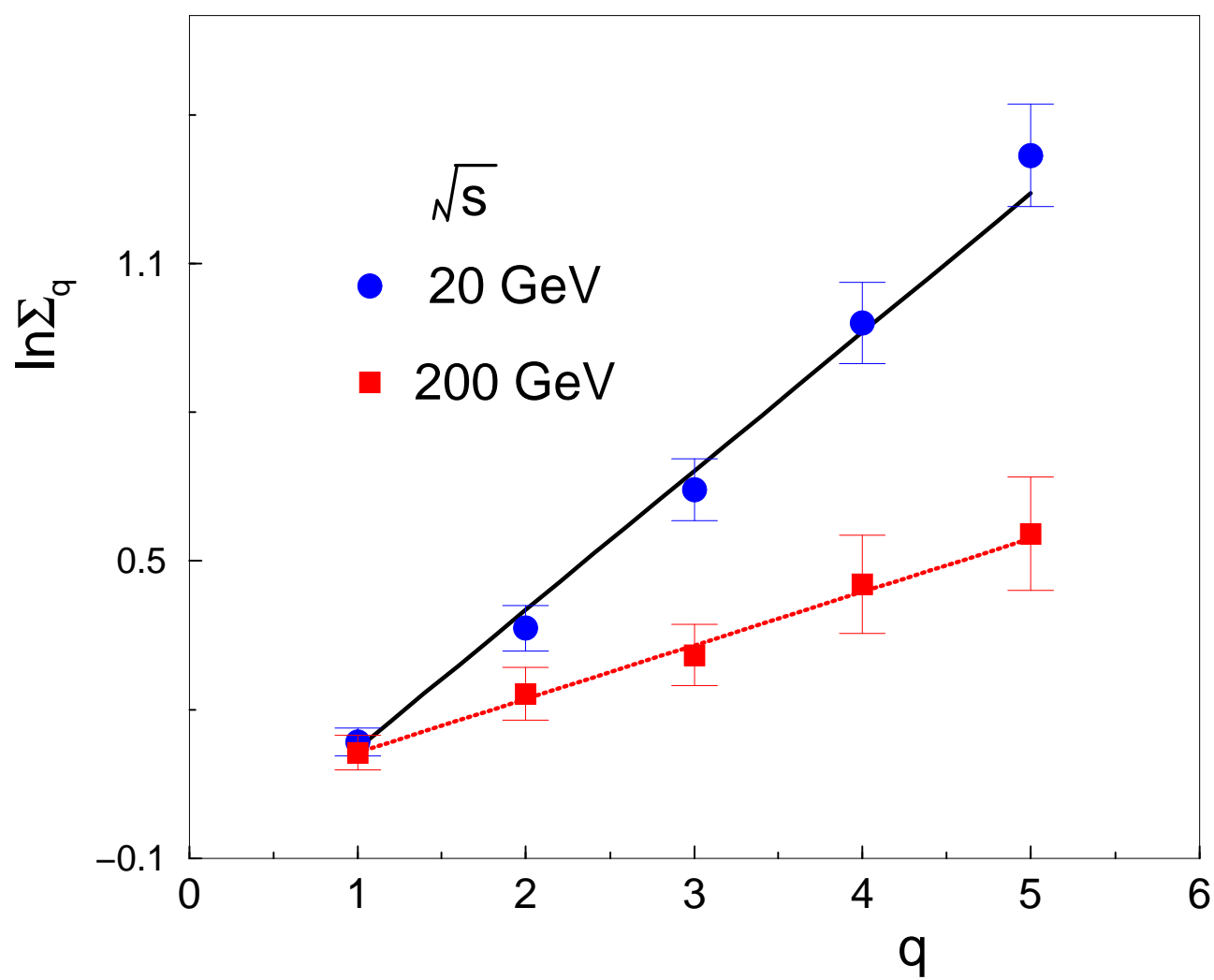


Fig.5